

10. Field Evaluation, Distribution and Maintenance

From "Verification, Validation, and Evaluation of Expert Systems, Volume I"

Evaluation, which includes field testing, addresses the issue "is the system valuable?" Value is indicated by the degree of end user approval, which in turn determines the extent of acceptance and use of the expert system. Distribution and maintenance of expert systems are addressed in this chapter.

Field Evaluation

Evaluation is the process of determining the likelihood that once deployed, the expert system will be used whenever appropriate. Evaluation should be an ongoing endeavor to help ensure maximum usage of the deployed system. Pertinent issues in evaluation are:

- Is the system user friendly and do the users accept the system?
- Does the system give "correct" results and is the logic of the system correct?
- Does the expert system offer an improvement over the practices it is intended to supplement?
- Is the system easy to learn and to become proficient on?
- Is the system useful as a training tool?
- Is the system in fact maintainable by other than the developers?
- Can the system be used in the intended work environment?

There are no universally accepted standards for the evaluation of expert systems. In fact it may be quite difficult to achieve. Sometimes evaluation is ignored until very late in system development. However, there are some things which can be done to make the process more effective. First, for systems under development, the developer should design for testing. For completed prototypes this is impossible, however workshops and substantial interaction with the target end users can make the process of field testing much more valuable. It is critical that the end users be kept aware of how important their contributions are and that their efforts are greatly appreciated.

Workshops and follow-up efforts with end users and experts can provide valuable improvements to an expert system and incentives for its use. After the expanded prototype has been constructed based on knowledge from the experts and end users on the development team, a workshop or series of workshops involving a larger community of experts and end users will usually result in major improvements to the system. The participants should be introduced to the computer program (expert system) and to the general concepts used in the development of the system. During the workshops the knowledge structure and the parameters used in the decision making process and their interrelationships should be reviewed. Expected benefits from workshops include:

- A verification check, i. e. review and improvement of the logic used in developing the system rules
- Enhancement of the knowledge base, i. e. finding oversights in the rules and the relationships between them
- More user oriented and user friendly interface
- Development of vested interest in the user community, i. e. establishment of a cadre of field users who want the system to succeed
- A better and more useable system

As part of the evaluation process, the assumptions made during planning of the system should be reexamined. For example during the planning process assumptions on the frequency of use, availability of input data, usefulness of system output, ease of use, etc. should have been documented. These assumptions should all be tested during field trials.

While there may be no universally accepted standards for field evaluation of expert systems, steps similar to those prescribed in Chapter 9, Testing, can improve the process. The big difference between testing and evaluation is that testing focuses on the accuracy of the advice given by the system while evaluation focuses on the degree of user acceptance of the system. The steps suggested for evaluation are as follows:

1. Select evaluation criteria and determine the minimum acceptable performance level for each criterion selected. Ideally these can be found in the requirements document for the expert system. If this information is missing (which is frequently the case) proposed criteria should be provided to the developers and the users of the system. For example:
 - a minimum of 90% of the users surveyed will agree that the consistency of answers the system provides offers a marked improvement over past practice and that the quality of answers provided is acceptable.
 - a minimum of 80% of users surveyed will rate the ease of learning and the user friendliness of the system as acceptable or highly acceptable.

- a minimum of 80% of users surveyed will rate the appropriateness of and ease of answering system queries as acceptable or highly acceptable.

A simple to use evaluation form should be prepared depicting these criteria.

2. Specify how each criterion will be measured. For example each of the criteria may be accessed by the users on the following three point scale:
 - highly acceptable
 - acceptable
 - unacceptable
3. Select a sample of users. Ideally a minimum of 12 users should be selected from the population of users. The greater the number of end users in the study and the more representative they are of the target end user community, the more the results can be viewed with confidence.
4. Gather data. Have each user to use the expert system on a minimum of six cases. After each case, have the user fill out the evaluation form, rating each criterion on the three point scale. Present the six cases to the users in random order. Debrief each user by asking for details on the basis of the rating given.
5. Analyze the data. Summarize all of the ratings of each user on each criterion and across all users for each criterion. Determine if the target level of acceptability has been reached or exceeded for each criterion. The data can then be analyzed.
6. Report the results and recommendations. Report on the strengths and weaknesses of the expert system. Concentrate on suggestions for improving the likelihood of user acceptance by emphasizing features which receive low ratings from the most users and those that could be the most quickly and economically improved.

The personnel who distribute the system and provide field support for the field evaluation must be carefully screened and selected. The wrong selection of support personnel can sabotage even the very best of expert systems. The field support personnel must have the following capabilities:

- Expert knowledge of the domain of the expert system. During this phase the field support staff may be called upon to not only provide support for the expert system that is dependent on domain knowledge, but to answer domain specific questions from the end users. A non domain expert can do irreparable damage to the credibility of the system during this phase.

- Sophisticated computer skills. The field support personnel have to go into a strange environment and correctly and efficiently install an expert system and then provide training using this unfamiliar equipment. Installation problems can always be expected.
- Excellent language skills in the language of the domain experts and end users. The field support personnel must be able to express concepts clearly and concisely to the users at their duty station and in their language and to understand the nuances that the end users try to convey. Anything less than FLUENCY in the language of the end users is not acceptable. In addition to language skills the field support personnel must have excellent interpersonal skills.

Some of the activities to be conducted in field testing are:

- Field operating environment - It is necessary to become acquainted with the field operating environment the expert system will be installed in. Even though operating environment was considered through the domain experts and end users, it will in fact appear different to every observer and there will be factors that effect operation that were not considered before the actual installation and field trials begin.
- Installation of expert systems - The expert system will have to be installed on the equipment provided by the end user/tester. This step will usually not be routine as computers or operating systems, etc., may have to be reconfigured to accommodate the expert system.
- Additional training for the end user/tester will need to be provided. Regardless of prior training, the end user/tester will need support to overcome the various nuances that appear during field test conditions. At this point it is also necessary to define the roles of various parties who participate in the field testing. Who runs the expert system? Who approves the use of expert system recommendations? Who applies the recommendations from the expert system? Who actually collects field data? Who does the preliminary screening/preparation of any data collected?
- After the expert system is installed on the end users computer the requirements and terms for the field tests must be reviewed. Competing demands on the end user are always more extensive at the end users duty station than they were during previous meetings. It is critical to get a renewed commitment. The commitment to support the end user/tester in every way possible must also be reaffirmed. The end user must understand how critical his/her support is and that the sponsor values this.
- Specific test cases should be identified and discussed. This includes previously identified sample test cases and new conditions that may be encountered in the field.
- The formal mechanism for incorporating findings from the field tests should be developed and reviewed in detail with the end user/tester. Also the formal mechanism for sharing

information between end users/testers must be developed and discussed with all end users/testers.

A final evaluation of end user satisfaction should take place after testing has demonstrated that the expert system has reached its target scope of coverage and level of accuracy.

Distributing and Maintaining Expert Systems

Once the expert system development effort has been completed, the tasks of distribution and maintenance begin. Although there are no fixed rules governing these tasks, there are some general guidelines which can make these tasks easier to perform.

Distribution

There are three major criteria that a developer should follow in order to facilitate the distribution of a given expert system.

1. Identify and involve the user community before starting the development of an expert system. This should insure that the expert system actually solves a real set of problems. Also it will give the user community a vested interest in the testing and application of the system.
2. Develop the system using standard hardware and software. Although there are a number of exotic pieces of expert systems hardware and software, the cost of these items is often high and the uncertainty of survival of these products in the market place makes it unreasonable to expect potential users to procure them just to use an expert system.
3. Use development software that does not require distribution licenses or where an unlimited distribution license can be purchased for a reasonable fee. The time and funding expended on paying fees for each system distributed, as required for many available development tools, can become an unwanted administrative and financial burden.

Maintenance

The task of system maintenance is one that should be planned for from the inception of the expert system development project. Maintenance can be facilitated by following a few good development rules. These include:

1. Design the expert system to be as transparent as possible. Since the system maintenance will probably not be conducted by system designers, it is necessary that the structure of the expert system be as straightforward and clear as possible.
2. The developers should use logical and understandable names for objects and knowledge structures within the system and avoid the use of cryptic names and obscure abbreviations.
3. The developers should also avoid the use of overly complex and obscure software structures, even though their use may provide some type of performance benefits.

4. Simplicity should be one of the guiding principals in the development effort.

The system must be well documented. The documentation should be produced as the system is developed, not as an after thought after the system is finished. The knowledge engineer should:

- Identify where the system's knowledge resides (e.g., in the knowledge base in the form of facts and rules and in the inference engine in the form of heuristic search techniques).
- Document the inference procedures that the system uses in producing solutions.
- Ensure that as part of the documentation an explicit model of the problem solver is included
- Ensure that the documentation also provides a comprehensive and well documented test procedure for the system

The expert system itself should contain an extensive set of both user "help" text and explanation text which explains how the system produced a given solution.

The documentation and the help and explanation text should be produced during the development phase and not added after the system has been built. One of the guiding principles that developers should use is "a poorly documented system will have a short useful life."

Each version of a given expert system should have a version number. This will make it easier to provide users with up-dated copies of the system.

Establish a mechanism for soliciting, receiving and acting upon feedback from the user community. This will facilitate the identification and removal of "bugs" in the system and will also make it easier to retrofit the system to satisfy specific user community needs after the system has been distributed to the user community.